PLAN DE INVESTIGACIÓN (PROYECTO DE TESIS DOCTORAL)

Behavioral Design of a Conversational AI to Support Unbiased Project Decision Making

Autor: Claudia Alcelay Larrarte

Director: Francisco José García Peñalvo

Plan de investigación

Programa de Doctorado: Formación en la Sociedad del Conocimiento

Universidad de Salamanca

1 de Junio, 2025

INTRODUCTION

Project managers are central to guiding projects towards successful completion, a process fundamentally driven by effective decision-making (1). However, the field of behavioral decision-making has extensively documented that actual human decision-making often significantly deviates from the rational ideal (2). These deviations are attributed to cognitive biases, which are systematic errors in thinking or reasoning. These biases result from using mental shortcuts, also known as heuristics (3). While heuristics can be adaptive in many situations, they can lead to maladaptive and erroneous decisions, termed biases, under certain conditions. Scholars have proposed that biases observed in controlled laboratory settings are also evident in real-world decision-making. Specifically, within project management, studies have begun to explore the influence of cognitive biases on managerial decisions. For instance, research has investigated the presence of anchoring and overconfidence biases in project managers' judgments. Furthermore, the very formulation of a project's aim is susceptible to heuristic thinking and the resulting cognitive biases, impacting the generation of alternative project goals and the assessment of their expected benefits. These biases can manifest in various stages of a project, including its initial conceptualization (4) and the communication of its scope to technical teams.

Artificial Intelligence (AI) is increasingly impacting and transforming decision-making processes across numerous domains, and project management is no exception (5). AI systems are being implemented to support a range of project activities, such as analyzing large datasets, automating repetitive tasks, and providing data-driven recommendations (6). Increasingly, conversational AIs like ChatGPT are also being used by project managers in everyday decision-making, offering real time support and content generation. Yet the development and deployment of these systems remain susceptible to human bias. However, the development and deployment of conversational AI systems are not immune to the influence of human biases. Such biases can be unintentionally encoded within the training datasets used to develop AI algorithms or introduced during the design and programming phases through developers' own biases and assumptions (5). As AI algorithms learn from data and identify patterns, they can also amplify existing biases present in the data (7).

A preliminary study conducted to justify the need of this research highlights a critical gap in the current understanding of how bias operates in AI-driven project management decision making. While the broader impact of AI on project decision-making has received growing attention (8), no systematic review has specifically addressed this issue of bias, which continues to shape the development of decision-support tools and applications.

This research carries significant potential benefits for a diverse group of stakeholders. Project managers will gain a more profound understanding of how cognitive biases can be amplified or newly introduced through the integration of AI tools into their decision-making processes, empowering them to adopt more critical evaluation practices and develop strategies for bias mitigation. AI developers can leverage the research findings to gain insights into the specific cognitive biases prevalent in project management, guiding them in the design and refinement of AI systems that are more robust, fair, and less likely to perpetuate biased outcomes. Organizational leaders can utilize the research to inform the implementation of targeted training programs focused on bias awareness and to develop informed policies and guidelines for the effective and responsible adoption of AI in project management contexts. Furthermore, policymakers can draw upon this research to inform the development of broader ethical guidelines and accountability frameworks for the application of AI within professional domains, ensuring fairness and transparency.

WORKING HYPOTHESIS AND PRINCIPAL OBJECTIVES SOUGHT

Hypothesis 1: AI impacts behavioral biases in project decision-making. This hypothesis assumes that conversational AI, particularly since the advent of ChatGPT, is significantly impacting behavioral biases within project decision-making processes. This shift underscores the importance of understanding the current landscape of how AI impacts these biases to provide a sound and solid base for future research.

Tasks:
- Conduct a systematic review of the academic literature and industry reports published since the advent of ChatGPT to assess the scope and nature of AI's impact on behavioral biases in project decision-making. (In progress)

Outcomes:
- Identification of frequently affected bias types; Mapping of decision contexts impacted by GenAI; Highlighting of project phases where bias is most likely to manifest; Identification gaps in research.

Hypothesis 2: Specific patterns define an AI's behavior and design, which can be addressed proactively during the AI development process. This hypothesis assumes that identifiable patterns in AI behavior and design influence decision-making biases. By recognizing these patterns early in the development process, AI systems can be designed to support unbiased decision-making.

Tasks:
- Investigate existing project management biases amplified by AI to identify recurring patterns or errors that contribute to biased decisions; Develop a set of guidelines for AI design that proactively addresses these patterns to minimize biases.

Outcome:
- A set of behavioral design guidelines for AI systems that emphasize bias mitigation through the example of one significative bias in project management.

Hypothesis 3: An AI model designed based on specific behavior patterns can effectively reduce biases in project management decision making. Building on the findings from Hypothesis 2, this hypothesis tests whether an AI model, designed with specific behavioral patterns in mind, can effectively reduce biases in setting and achieving project objectives.

Tasks:
- Develop a multi-agent AI system incorporating the guidelines established in Hypothesis 2.
- Implement and test the AI system in a controlled environment to evaluate its impact on reducing biases in project management scenarios.

Outcomes:
- A pilot product to be tested
- Validation of the behavioral design guidelines through empirical evidence.

METHODOLOGY

Application of the Design Science Research Methodology (DSRM).

This research applies the Design Science Research Methodology (DSRM) as outlined by Peffers et al. (9) to explore and address cognitive biases in Conversational AI-supported project management decision-making. The approach ensures practical outcomes grounded in theory and academic rigor. Each step is adapted to the research timeline and scope.

Step 1: Problem Identification and Motivation (2024–2025). Systematic review.
The study addresses the growing concern of how Generative AI, especially since the launch of ChatGPT in late 2022, may introduce or amplify cognitive biases in project decision-making. A preliminary review reveals a lack of systematic studies in this area. The high impact of project decisions and increasing use of AI tools make this a timely and critical issue.

To investigate it rigorously, a systematic literature review is being conducted across Scopus and ProQuest databases. The research questions have been initially formulated using the PICOC framework and are currently being refined in light of recent developments and publications (10). Expected outcomes of this systematic review are:

- Identification of frequently affected bias types (e.g., anchoring, framing, automation bias)
- Mapping of decision contexts impacted by GenAI (e.g., risk, prioritization, resource trade-offs)
- Highlights project phases where bias is most likely to manifest
- Identification of gaps in existing conceptual and empirical research

Step 2: Define Objectives (2024–2025)
In parallel with the systematic literature review conducted in Step 1, this stage defines the research objectives, ensuring that they reflect the types of decisions most frequently impacted by Conversational AI in project settings. To guide this analysis, the review adopts a principle-based coding framework derived from the PMBOK Guide – Seventh Edition.

This framework focuses on six principles identified as particularly vulnerable to the influence of Conversational AI due to their behavioral orientation and proximity to judgment and bias: leadership, stakeholders, risk responses, value, context, and systems.

These principles serve as a lens to structure the analysis and formulate research objectives that are directly connected to project decision-making dynamics under AI influence.

Step 3: Design and Development (2025–2026)
A multi-agent AI system will be developed based on behavioral design principles. It will help project managers make decisions while minimizing biases like anchoring and overconfidence, grounded in the first two hypotheses.

Step 4: Demonstration (2026–2027)
The artifact will be tested in a simulated project setting to show how it supports unbiased decision-making through interactive scenarios.

Step 5: Evaluation (2026–2027)
Effectiveness will be measured via bias reduction, decision quality, and user feedback, using both qualitative and quantitative methods. Findings will inform iterative refinements.

Step 6: Communication (2026–2027 and ongoing)
Results will be shared through conferences (beginning with TEEM in 2024), journal articles, and academic collaborations.

The final thesis will contribute to the fields of project management, AI ethics, and behavioral decision-making.

MEDIOS Y RECURSOS MATERIALES DISPONIBLES (MÁXIMO 50 LÍNEAS):
MATERIAL MEANS AND RESOURCES AVAILABLE (50 LINE MAXIMUM):

I am currently enrolled in the Doctoral Program in Education in the Knowledge Society at the University of Salamanca, which provides access to a range of academic resources, including Scopus and Web of Science. I also take part in the doctoral training activities offered through the program and use the university's software and academic tools as part of my research process.

As part of this process, I regularly consult the program's online portal (https://knowledgesociety.usal.es), which hosts a range of materials on interdisciplinary research in the context of the Knowledge Society. Publications available through the portal—such as García-Peñalvo (2014) and García-Peñalvo et al. (2019)—have supported my understanding of the theoretical and methodological underpinnings of the program, particularly regarding educational innovation and doctoral training models (11, 12).

In addition, I have engaged with the work of the GRIAL Research Group (https://grial.usal.es) to inform the design and conceptual structure of my thesis. The group's research on technology-enhanced learning, digital transformation, and educational systems provides useful frameworks that align with the objectives of my study. The description offered in García-Peñalvo et al. (2019) has been particularly helpful in understanding the group's main lines of inquiry and their relevance to current educational challenges (13).

TIMING SCHEDULE OVER FOUR YEARS

YEAR 1. 2024-2025. Systematic Literature Review

I am currently conducting a systematic literature review using the Scopus and ProQuest databases. Following the PICOC framework, I have defined the research questions, scope of the review, expected outcomes, and inclusion and exclusion criteria. I am now in the second screening phase, during which the research questions are being refined and adjusted as needed.

In parallel, I am actively participating in the doctoral training sessions offered by the university. I plan to present preliminary results of my research at the TEEM Congress in October and submit a full article for publication in November.

YEAR 2. 2025-2026

Investigate existing project management biases amplified by AI to identify recurring patterns or errors that contribute to biased decisions; Develop a set of guidelines for AI design that proactively addresses these patterns to minimize biases.

In parallel, I will be actively participating in the doctoral training sessions offered by the university. I also plan to present preliminary results of my research at various academic conferences and to submit a second article for publication in November.

YEAR 3. 2026-2027

I will design and test an artifact intended to address one or more specific instances of the bias related identified problems. Assess and measure how the artifact resolves or mitigates the bias problem. Compare the initial objectives with the observed results and evaluate the outcomes.

In parallel, I will be actively participating in the doctoral training sessions offered by the university. I also plan to present preliminary results of my research at various academic conferences and to submit a second article for publication in November.

PERSONAL TRAINING PLAN

My personal training plan includes attendance and accreditation of the completeness of:

- Recursos de información electrónicos de la USAL
- Fuentes especializadas: Bases de datos (EBSCO, PROQUEST, INDICES-CSIC, SCIELO)
- Fuentes especializadas: Índices de citas (WoS, Scopus, Google Académico)
- Gestores bibliográficos y Normas de redacción científica: Refworks, Zotero, Mendeley
- Dónde publicar: Indicadores de calidad y acuerdos transformativos
- Identidad digital o cómo mejorar tu perfil investigador
- Búsquedas bibliográficas, mapeo sistemático y revisión sistemática de la literatura

- I have registered for a training course on Open Science, Open Access Research, and Citizen Science, offered as part of the University's transversal doctoral training program, to be provided in Summer 2025. I am currently awaiting confirmation of acceptance.

REFERENCES

(1) Flyvbjerg, B. (2021). Top ten behavioral biases in project management: An overview. Project Management Journal, 52(6), 531–546. https://doi.org/10.1177/87569728211049046

(2)Oliveira, L., & Figueiredo, P. S. (2020). An exploratory study on the presence of cognitive biases in project management decisions at a research institute. Brazilian Journal of Development, 6(9), 68693–68707. https://doi.org/10.34117/bjdv6n9-349

(3) Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. Annual Review of Psychology, 62, 451–482. https://doi.org/10.1146/annurev-psych-120709-145346

(4) Kiełczewski, D., Matel, A., & Poskrobko, T. (2016). Project management in behavioural perspective – Cognitive biases in the formulation of the aim of the project. Engineering Management in Production and Services, 8(3), 70–78. https://doi.org/10.1515/emj-2016-0025

(5) Soleimani, M., Intezari, A., Taskin, N., & Pauleen, D. J. (2021). Cognitive biases in developing biased artificial intelligence recruitment system. In Proceedings of the 54th Hawaii International Conference on System Sciences (pp. 620–629). https://doi.org/10.24251/HICSS.2021.620

(6) Rizzo, C., Bagna, G., & Tuček, D. (2024). Do managers trust AI? An exploratory research based on social comparison theory. Management Decision. https://doi.org/10.1108/MD-06-2023-0837

(7) Soleimani, M. (2022). Mitigating cognitive biases in developing AI-assisted recruitment systems: A knowledge-sharing approach. International Journal of Knowledge Management, 18(1), 1–18. https://doi.org/10.4018/IJKM.296260

(8) Varsha, P. S. (2023). How can we manage biases in artificial intelligence systems: A systematic literature review. International Journal of Information Management Data Insights, 3(1), 100165. https://doi.org/10.1016/j.jjimei.2023.100165

(9) Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. Journal of Management Information Systems, 24(3), 45–77. https://doi.org/10.2753/MIS0742-1222240302

(10) García-Peñalvo, F. J. (2022). Developing robust state-of-the-art reports: Systematic Literature Reviews. Education in the Knowledge Society, 23, Article e28600. https://doi.org/10.14201/eks.28600

(11) García-Peñalvo, F. J. (2014). Formación en la sociedad del conocimiento, un programa de doctorado con una perspectiva interdisciplinar. Education in the Knowledge Society, 15(1), 4-9. https://doi.org/10.14201/eks.11641

(12) García-Peñalvo, F. J., Rodríguez-Conde, M. J., Verdugo-Castro, S., & García-Holgado, A. (2019). Portal del Programa de Doctorado Formación en la Sociedad del Conocimiento. Reconocida con el I Premio de Buena Práctica en Calidad en la modalidad de Gestión. In A. Durán Ayago, N. Franco Pardo, & C. Frade Martínez (Eds.), Buenas Prácticas en Calidad de la Universidad de Salamanca: Recopilación de las I Jornadas. REPOSITORIO DE BUENAS PRÁCTICAS (Recibidas desde marzo a septiembre de 2019) (pp. 39-40). Ediciones Universidad de Salamanca. https://doi.org/10.14201/0AQ02843940

(13) García-Peñalvo, F. J., Rodríguez-Conde, M. J., Therón, R., García-Holgado, A., Martínez-Abad, F., & Benito-Santos, A. (2019). Grupo GRIAL. IE Comunicaciones. Revista Iberoamericana de Informática Educativa(30), 33-48. https://bit.ly/35IIQh9